Mathematical Statistics

Support - set of outcomes with positive probability

Bernoulli Distribution: $[0,1]$ with $[1-p, p]$
$\mathrm{E}[\text{Bernoulli}] = p$
$\mathrm{Var}[\text{Bernoulli}] = p(1-p)$
$P_X(s) = ps + q$
$m_Y(t) = (1-p) + pe^t$

Binomial Distribution (n,p):
$[0,1,\dots,n]$ with $[\binom{n}{0}p^0(1-p)^n, \dots]$
$\mathrm{E}[\text{Binomial}(n,p)] = np$
$\mathrm{Var}[\text{Binomial}(n,p)] = np(1-p)$
$P_X(s) = (ps+q)^n$
$m_Y(t) = (pe^t + (1-p))^n$

Geometric Distribution: $[0,1,2,\dots]$ with $[p, p(1-p), p(1-p)^2, \dots]$
$\mathrm{E}[\text{Geometric}] = \frac{1-p}{p}$
$\mathrm{Var}[\text{Geometric}] = \frac{1-p}{p^2}$
$P_X(s) = \frac{p}{1-qs}$
$m_Y(t) = \frac{p}{1-(1-p)e^t}$

Poisson Distribution: $[0,1,2,\dots]$ with $[e^{-\lambda}\frac{\lambda^k}{k!}]$
$\mathrm{E}[\text{Poisson}] = \lambda$
$\mathrm{Var}[\text{Poisson}] = \lambda$
$P_X(s) = e^{\lambda(s-1)}$
$m_Y(t) = e^{\lambda(e^t-1)}$

Uniform Distribution $[a,b]$:
$f_Y(y) = \frac{1}{b-a}1_{[a,b]}(y)$
$\mathrm{E}[Y] = \frac{a+b}{2}$
$\mathrm{Var}[Y] = \frac{(b-a)^2}{12}$
$F_Y(y) = \frac{y-a}{b-a}$ for $y \in [a,b]$
$m_Y(t) = \frac{e^{bt}-e^{at}}{t(b-a)}$
$\mu_k = \frac{b^{k+1}-a^{k+1}}{(k+1)(b-a)}$

Normal Distribution $Y \sim N(\mu, \sigma)$:
$y\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(y-\mu)^2}{2\sigma^2}}$
$\mathrm{E}[Y] = \mu$
$\mathrm{Var}[Y] = \sigma^2$
$\mu_k^c = \sigma^k(k-1)(k-3)\dots(1)$ and $\mu_k^c = 0$ for odd k
$m_Y(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

Exponential Distribution $\tau > 0$:
$f_Y(y) = \frac{1}{\tau}e^{-\frac{y}{\tau}}1_{[0,\infty)}(y)$
$\mathrm{E}[Y] = \tau$
$\mathrm{Var}[Y] = \tau^2$
$\mu_k = k!t^k$
$S(y) = e^{-\frac{y}{\tau}}$
$h(y) = \frac{1}{\tau}$
$F_Y(y) = 1 - e^{-\frac{y}{\tau}}$ for $y > 0$
$m_Y(t) = \frac{1}{1-\tau t}$

$\chi^2(n)$ Distribution:
$f_Y(y) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}y^{\frac{n}{2}-1}e^{-\frac{y}{2}}$
$\mathrm{E}[Y] = n$
$\mathrm{Var}[Y] = 2n$
$F_Y(y) = \frac{1}{\Gamma(\frac{n}{2})}\gamma(\frac{n}{2}, \frac{y}{2})$
$m_Y(t) = (1-2t)^{-\frac{n}{2}}$

$\Gamma(k,\tau)$ Gamma Distribution:
$f_Y(y) = \frac{1}{\Gamma(k)\tau^k}y^{k-1}e^{-\frac{y}{\tau}}$
$\mathrm{E}[Y] = k\tau$
$\mathrm{Var}[Y] = k\tau^2$
Exponential $E(\tau) = \Gamma(1,\tau)$ and $\chi^2(n) = \Gamma(\frac{n}{2}, 2)$
$m_Y(t) = (1-\tau t)^{-k}$

$\mathrm{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y)dy$ when well defined

$\mathrm{Var}[Y] = \int_{-\infty}^{\infty}(y-\mu_y)^2 f_Y(y)dy$ with $\mu_Y = \mathrm{E}[Y]$ mean/expectation of Y

$\text{Cov}(X, Y) = \text{E}[(X - \text{E}[X])(Y - \text{E}[Y])] = \text{E}[XY] - \text{E}[X]\text{E}[Y]$

k-th Moment (Raw):
$\mu_k = \text{E}[Y^k] = \int_{-\infty}^{\infty} y^k f_Y(y)dy$

k-th Central Moment: $\mu_k^c = \text{E}[(Y - \text{E}[Y])^k] = \int_{-\infty}^{\infty} (y - \mu)^k f_Y(y)dy$

Note Standardized Moment is Central Moment normalized typically with division by an expression of the Variance which renders the moment scale invariant.

Expectation/Mean $\mu = \mu_1 = \text{E}[Y]$

Variance $\mu_2^c = \text{Var}[Y]$

Skewness $\text{E}[\frac{Y - \text{E}[Y]}{sd[Y]}^3] = \frac{\mu_3^c}{(\mu_2^c)^{\frac{3}{2}}}$

Kurtosis $\text{E}[\frac{Y - \text{E}[Y]}{sd[Y]}^4] = \frac{\mu_4^c}{(\mu_2^c)^2}$

Cumulative Distribution Function [cdf]:
$F(y) = P[Y \leq y]$

$F(y) \int_{-\infty}^{y} f(z)dz$

$f(y) = F'(y)$

Survival Function: $S(y) = 1 - F(y)$

Hazard Function: $h(y) = \frac{f(y)}{S(y)}$ roughly the conditional probability that the individual will die at time y given that it has survived until y

cdf-Method: $W = g(Y)$ want
$F_W(w) = P[g(Y) \leq w] = P[Y \leq g^{-1}(w)] = F(g^{-1}(w))$

Inverse Exponential Distribution:
$f(y) = \frac{1}{ty^2} e^{-\frac{1}{ty}} 1_{(0,\infty)}(y)$

$F(y) = e^{-\frac{1}{ty}}$

χ^2 Distribution:
$f(y) = \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}} 1_{(0,\infty)}(y)$

$f_W(w) = f_Y(g^{-1}(w))|(g^{-1})'(w)|$

$m_Y(t) = \text{E}[e^{tY}] = \int_{\infty}^{\infty} e^{ty} f_Y(y)dy$

$W = aY + b$, $m_W(t) = e^{tb} m_Y(at)$

$Y_1, Y_2, \ldots, Y_n$ independent
$Y = Y_1 + Y_2 + \cdots + Y_n$ then
$m_Y(t) = m_{Y_1}(t) \cdot m_{Y_2}(t) \cdots m_{Y_n}(t)$

$m_Y(t) = \sum_{k=0}^{\infty} \frac{\mu_k}{k!} t^k$, $\mu_k = \text{E}[y^k]$ is k-th moment of Y

$\mu_k = \frac{d^k}{dt^k} m_Y(0)$

Central Limit Theorem: $Y_1, Y_2, \ldots, Y_n$ independent random variables with the same distribution. If $\mu = \text{E}[Y_i]$, $\text{Var}[Y_i] < \infty$ then the distribution of the normalized sum $\frac{S_n - \text{E}[S_n]}{sd[S_n]}$

If $np > 10$ and $n(1 - p) > 10$ then Normal Approximation

If $n > 50$ and $np < 5$ then Poisson $\lambda = np$

χ^2-Distribution with n degrees of freedom is the distribution of a sum $W = Z_1^2 + Z_2^2 + \cdots + Z_n^2$ of squares of n independent unit normal $N(0, 1)$ random variables denoted by $\chi^2(n)$

Estimator: function of data which does not depend on the value of unknown parameters i.e. is based on the sample. Often denoted with a hat, for example

$\hat{\mu} = \frac{Y_1+Y_2+\cdots+Y_n}{n}$ the sample mean may be used to estimate the mean.

$\text{Bias}(\hat{\theta}) = \text{E}[\hat{\theta}] - \theta$ if 0 unbiased

Sampling distribution of $\hat{\mu}$ is $N(\mu, \frac{\sigma}{\sqrt{n}}$

$\hat{Y} = \frac{Y_1+\cdots+Y_n}{n}$ Sample Mean

$\frac{1}{n}\sum_{k=1}^{n}(Y_k - \mu)^2$ Sample Variance Known Mean

$\frac{1}{n-1}\sum_{k=1}^{n}(Y_k - \hat{Y})^2$ Sample Variance Unknown Mean

Error of $\hat{\theta}$ is $\hat{\theta} - \theta$

Absolute Error of $\hat{\theta}$ is $|\hat{\theta} - \theta|$

Relative Error of $\hat{\theta}$ is $|\frac{\hat{\theta}-\theta}{\theta}|$

Squared Error of $\hat{\theta}$ is $(\hat{\theta} - \theta)^2$

Mean-Squared Error of $\hat{\theta}$ is $MSE(\hat{\theta}) = \text{E}[(\hat{\theta} - \theta)^2] = \text{Var}[\hat{\theta}] + (\text{Bias}(\hat{\theta}))^2$

Standard Error of $\hat{\theta}$ is $se(\hat{\theta}) = \sqrt{\text{Var}(\hat{\theta})}$

Pivotal Quantity function of sample data and parameter θ whose distribution does not depend on θ.

For example $\hat{Y} - \mu$ from $N(\mu, 1)$ is a pivotal quantity normally distributed with Mean 0 and Variance $\frac{1}{n}$

Likelihood Function $L(\theta : y_1, \ldots, y_n) = f^{\theta}(y_1) \cdot f^{\theta}(y_2) \ldots f^{\theta}(y_n)$ is the pdf for this sample from the distribution given by θ

To compute posterior update prior via this.

Maximum Likelihood Estimator maximizes L for a sample.

Log likelihood function can be helpful to take the derivative of a sum and compute extremum.

Bayes Estimator and Credible Interval

Beta Distribution $\text{Beta}(\alpha, \beta)$:
$f(y) = \frac{1}{B(\alpha,\beta)}y^{\alpha-1}(1-y)^{\beta-1}$
$\text{E}[Y] = \frac{\alpha}{\alpha+\beta}$
$\text{Var}[Y] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

Sufficient if and only if $L(\theta, y_i) = g(\theta, T(y_i))h(y_i)$